

Systematic Validation and Improvement of Quantum Chemistry Methods for the Prediction of Physical and Chemical Properties

The research described here will lead to the systematic validation and improvement of quantum chemistry techniques for the prediction of physical and chemical properties of a large variety of chemical systems. The primary goal is to provide uncertainties in properties computed by the most popular levels of theory making use of standard statistical techniques and high-quality experimental data. This work is also focused in making available procedures and data that will enable researchers in industry to choose the most reliable quantum chemistry methods for prediction of a physical and/or chemical property. An in-depth understanding of the sources of uncertainties also will lead to improvements of the quantum chemistry methods used in the predictions

C.A. Gonzalez, R. D. Johnson, and K.K. Irikura
(Div. 838)

The use of modern computational chemistry methodologies in the prediction of molecular properties has become increasingly popular mainly due to significant improvements in the algorithms, the accuracy of the methods, and the advent of powerful computer resources. This is particularly true in the area of thermochemistry, where researchers in industry and academia perform quantum chemistry calculations on a routine basis. Recently, there has been an increased interest in the use of theoretical predictions of physical and chemical properties made by state-of-the-art quantum chemistry methodologies to fill gaps in the available experimental data. This approach could potentially eliminate the cost of experimental measurement and will be particularly useful in cases where extrapolations based on available experimental data are not possible or are unreliable. However, despite the aforementioned progress, the predictive power of most of the quantum chemistry methodologies has not been established on solid ground due to the lack of knowledge of their relative uncertainties. To understand these uncertainties an exhaustive and systematic validation of these methodologies involving robust strategies to compare theory and experiment is necessary. In addition, the results of such validations must be widely available to the scientific community to provide a basis for new scientific discoveries.

NIST researchers are working on an integrated approach that will lead to the systematic validation as well as improvements of popular and new theoretical methods using high-quality experimental data. This approach involves the development of state-of-the-art databases in order to disseminate the accuracy, reliability, and

transferability of the most popular quantum chemistry methodologies in the prediction of physical and chemical properties of a large variety of chemical systems ranging from isolated atoms and molecules to condensed phases. The theoretical methodologies being studied include semi-empirical Hamiltonians, wavefunction-based *ab initio* molecular orbital theory, and Density Functional Theory (DFT) while the experimental data used in the validation include measurements already available in the literature and also made by researchers at NIST.

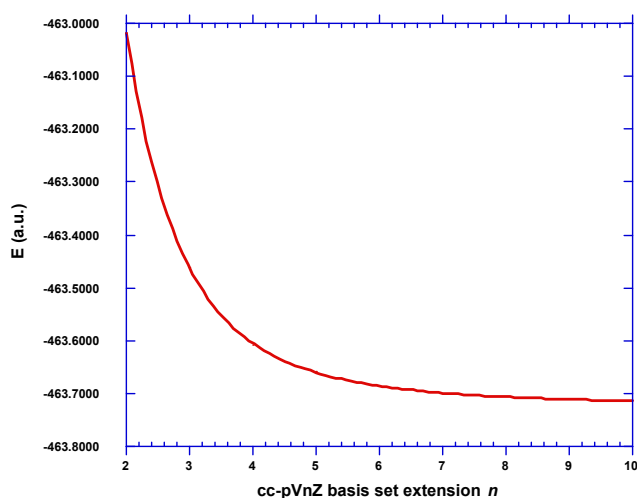
Assessing the Accuracy and Reliability of Møller-Plesset Perturbation Theory in Describing Dispersion Forces in van der Waals Clusters of Aromatic Molecules.

Aromatic-aromatic interactions play an important role in many chemical and biological systems. They control, among others, the base-base interactions leading to the double helical structure of DNA, the function of the special pair in photosynthetic reaction centers, the packing of aromatic crystals, the formation of aggregates, the binding affinities in host-guest chemistry, and the conformational preference of polyaromatic macrocycles and chain molecules. As the species formed due to intermolecular interactions, the geometrical structures of aromatic dimers and higher clusters, generated by free jet expansion, provide important information concerning the nature of the forces between aromatic molecules and other clusters properties. This information is highly useful in a wide variety of applications such as photochemistry, catalysis, homogeneous nucleation, the structure of condensed matter, and fabrication of nanodevices. For these reasons, the structural probe of the small clusters of aromatic hydrocarbons has been the subject of considerable interest in recent years. Unfortunately, because clusters are bound by weak electrostatic and van der Waals (vdW) forces, they tend to have floppy structures that are difficult to characterize experimentally. Quantum chemistry calculations are therefore useful for interpreting experimental observations and for making structural predictions that could guide experimental studies. Reliable *ab initio* calculations of aromatic clusters must include electron correlation explicitly in order to obtain good representations of dispersion and electrostatic forces that are responsible for the binding of the species. The most popular method to compute the interaction energies D_e in vdW clusters is based on the supermolecular approach, where D_e is calculated as the difference between the total energy of the vdW complex and the total energy of the separated monomers. Most of the time the total energies are computed by the *ab initio* second order Møller-Plesset perturbation theory, MP2, which provides a good compromise between computational expense and

accuracy. Although in principle MP2 is capable of describing dispersion forces, currently, there is not clear measure of the accuracy and/or reliability of the method when treating vdW clusters of aromatic molecules.

NIST researchers have designed a simple diagnostics in order to test the validity of MP2 in the case of benzene dimers. This procedure was used to quantify the intrinsic error (no basis set dependence) in the MP2 methodology as well as the errors arising from the limitation in the basis sets employed. The results of this work indicate that in the case of benzene dimers, MP2 predicts a binding energy 100% larger than the corresponding experimental value! In addition, it is also shown that the errors obtained when using relatively small (and in principle less accurate) basis sets cancel the intrinsic error in MP2 in a fortuitous manner, providing a rational explanation for the excellent agreement observed between the binding energies computed with MP2 using small basis sets such as 6-31G and the experimental results obtained in the case of larger aromatic vdW clusters.

Energy vs Basis Set Extension n Computed with MP2/cc-pVnZ Basis for Benzene Dimer in Parallel Displaced Conformation

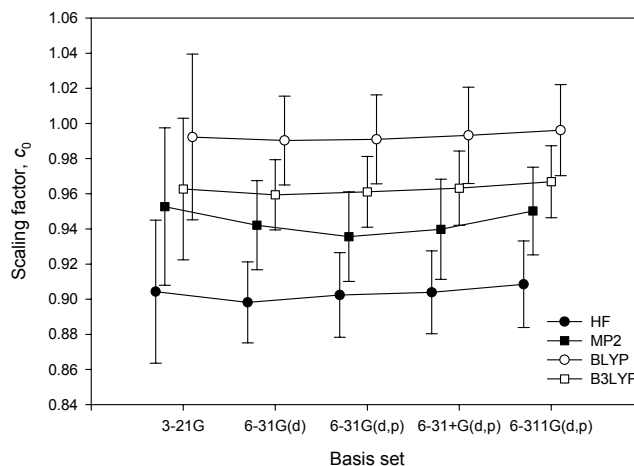


Virtual Measurements: Quantifying Uncertainty in Theoretical Predictions.

It is well recognized that a complete expression of experimental measurement includes both the central value and its associated uncertainty. The central value alone is incomplete. Throughout computational science, the results of models are usually reported without their associated uncertainties, making them incomplete. The term *virtual measurement* refers to a prediction from a computational model together with its associated uncertainty, and emphasizes the analogy with experimental (physical) measurement. Our investigation of the uncertainties associated with predictions from quantum chemistry models are intended to help “virtual measurements” supplant “calculated results.” The results from *ab initio* calculations differ from experiment by systematic biases, not random errors. For a given *ab initio* calculation, corrections can be applied to compensate for this bias and an uncertainty can be determined

from the distribution of corrections for a class of molecules. We use the experimental and *ab initio* data contained in the NIST Computational Chemistry Comparison and Benchmark Database (CCCBDB), which contains data for over 600 molecules, over 100 calculations for each molecule, and over 4000 vibrational frequencies. We combine this data with analysis based upon the *Guide to the Expression of Uncertainty in Measurement*, published by the International Organization for Standardization (ISO), to determine biases and uncertainties. We have applied this to enthalpies of formation (ref. 1) and to vibrational frequencies (ref. 2). This allows us to upgrade these “calculated values” into “virtual measurements”. For the enthalpies of formation we found the class of molecules to be critical in determining the bias and uncertainties. For the vibrational frequencies we found the uncertainties to be orders of magnitude larger than previously believed. As a bonus, trends were extracted and used to provide general recommendations for estimating scaling factors for arbitrary combinations of theory and basis set.

Vibrational Scaling Factors



Growth of the CCCBDB.

The Computational Chemistry Comparison and Benchmark Database (CCCBDB) is a collection of experimental and calculated *ab initio* thermochemical, vibrational (mentioned above), geometric and electrostatic data for 680 molecules and a web site of over 400 pages to display and compare the experimental and computed properties. The database contains results from over 107,000 calculations. A major upgrade in the database software, from Microsoft Access to Microsoft SQL Server*, should improve the efficiency and reliability of data transfer to users. Many small improvements have been made to the look and feel to provide easier access and navigation through the web-site. New tools include the ability to view animated vibrational modes and interactive pages for displaying bond lengths and angles. New classes of chemistry are represented by the addition of molecules that contain bromine, titanium, copper, or zinc, more aromatic species, and hydrogen-bonded systems.



<http://srdata.nist.gov/cccbdb/default.htm>

References:

- 1) Irikura, K. K.; Johnson, R. D., III; Kacker, R. N. *Metrologia* **41**, 369-375, 2004.
- 2) Irikura, K. K.; Johnson, R. D., III; Kacker, R. N. *J. Phys. Chem. A* **2005**, *109*, 8430-8437.

*Disclaimer:

Certain commercial equipment, instruments, or materials are identified in this document. Such identification does not imply recommendation or endorsement by the [National Institute of Standards and Technology](#), nor does it imply that the products identified are necessarily the best available for the purpose.

Future Plans: In the near future, this work will be extended to validate other quantum chemistry methodologies including DFT, wavefunction-based *ab initio* molecular orbital theory, and semi-empirical Hamiltonians. The virtual measurements techniques will be applied to vibration-dependent thermochemical properties such as zero-point energies, heat capacities, and entropies. As part of that work, it will be necessary to prepare some new, evaluated databases of benchmark experimental measurements. Also, performance of these quantum chemistry methodologies in describing physical and chemical properties of metals and organometallic systems (important in material science, biochemical processes, nanotechnology, as well as in catalysis) will also be assessed.